Economics and Environment – Modeling Global Linkages

Bouwmeester, Maaike C.

Chapter 3

Specification and aggregation errors in environmentally extended, international input–output models[1,2]

3.1 Introduction and literature review

In recent decades, environmental concerns continued to grow, leading to mitigation policies at all levels of governance. For some environmental issues, such as global climate change due to greenhouse gas emissions, the optimal scope is clearly the global level. Other issues result in more localized problems, such as water shortages, but their cause may also be global because of the growing importance of global value chains. Indicators monitoring environmental impacts, therefore, should account explicitly for the globalization of production and be based on international comparative data that acknowledge the importance of global value chains. The most detailed data to date that satisfy these requirements are found in the environmentally extended international input–output (EE-IIO) table that is assembled in the EU-funded EXIOPOL project (Tukker et al., 2009), which we will use in this article.

---

[1] Published as Bouwmeester and Oosterhaven (2013).

[2] This work is part of the EXIOPOL project (http://www.feem-project.net/exiopol), an integrated project funded by the European Union, under Framework Programme 6, Priority 6.3 Global Change and Ecosystems, Grant agreement no. 037033-2. We thank our partners for their cooperation and Arnold Tukker, the project leader, for stimulating comments. We also thank Richard Wood, Glen Peters, an anonymous referee, and the editor Christian Vossler, for their their comments on earlier versions of the chapter, and Elisabeth Nevins Caswell for editing out English.

Both in the policy debate and in the scientific literature the attention has shifted from measuring emissions and resource use at producer locations to attributing them to consumer locations (see Wiedmann, 2009). At the national level, the latter requires emissions and resource use embodied in imports to be included in the estimates, whereas those embodied in exports need to be excluded (Peters, 2008; Peters and Hertwich, 2008; Serrano and Dietzenbacher, 2010, Davis and Caldeira, 2010; Kanemoto et al., 2012).

Including emissions and resource use embodied in imports requires information about the production technologies of trading partners, their emissions, and their resource use. An EE-IIO table provides such information (Wiedmann et al., 2007). Users of single-country input–output models, however, often assume that the environmental and interindustry coefficients of the country at hand resemble the coefficients of the trading partners. This domestic technology assumption (DTA) enables the use of single-country data to estimate emissions embodied in both its imports and its exports. When technologies differ, the DTA creates specification errors. Inaccurate estimates may result in ineffective environmental policies when environmental impacts are underestimated or they may divert the focus from other environmental problems when they are overstated.

Peters, Andrew and Lennox (2011) compare various models to estimate CO_2 emissions of consumption, including a single-country model (with the DTA), and a full multi-regional input–output (MRIO) model. To build the latter, these authors use GTAP data for 2004, which contain 113 regions and 57 commodities. The GTAP data are balanced by forcing the IO tables to match macroeconomic and reconciled trade data (Narayanan and Walmsley, 2008). Peters, Andrew and Lennox (2011) add a proportionality assumption to distribute bilateral trade over the consuming industries. Comparing the DTA estimates against the MRIO estimates reveals substantial deviations; the former results in higher values. Peters, Andrew and Lennox (2011) recommend that researchers should not adopt the domestic technology assumption without strong justification.

In this article, we develop a methodology that presents an additive decomposition of the error resulting from the DTA assumption into its constituent elements. The overall error is decomposed into the error due to assuming that foreign industries have domestic interindustry coefficients and the error due to assuming that foreign industries have domestic emission coefficients. We apply the methodology to embodied CO_2 emissions and embodied water use, as calculated with the EXIOPOL EE-IIO table. Our results show that the total DTA specification error is unacceptably large, despite that the domestic emission error is partially cancelled out by the domestic interindustry error, especially for CO_2 emissions.

The current state of MRIO data sets, as described by Wiedmann et al. (2011), suggest an improvement over previously used data, though they still contain aggregated industries and approximations of missing information. The resultant errors can be studied using the newer generation of environmentally extended IO models. An early example is provided by Lenzen et al. (2004), who study international feedback effects and the errors associated with a single-region IO model compared with a model with four countries and a 'Rest of the World' region. This unidirectional trade model includes the technologies of trading partners but only through trade linkages with the country at hand. They also study the effect of sector aggregation; their findings suggest that it is important to include a trading partner's interindustry and emission coefficients explicitly. Large errors result when they aggregate industries into just ten sectors.

Andrew et al. (2009) show that with just a few regions, a unidirectional trade model can give a reasonable approximation of the full model. Bilateral linkages between other countries either are excluded or approximated through a DTA. Including the trade partner that is responsible for the largest share of embodied emissions also can significantly improve the estimates. A more extensive method to study the emissions embodied in imports is provided by the 'emissions embodied in bilateral trade' (EEBT) approach of Peters (2008). When applying EEBT, emissions embodied in imports of country *r* are estimated by calculating the emission embodied in all exports to *r*. The data requirements for this approach are extensive, as it requires IO tables for all countries considered as well as for their trading partners, and trade data to establish the bilateral trade flows. However, the approach does not allow the calculation of indirect international feedback effects, which is possible in MRIO modeling. Su and Ang (2011) find significant differences between estimates resulting from the EEBT and MRIO models.

Combining two data sets often requires some aggregation or disaggregation of one of the data sets. When the IO data provide more sector detail than the environmental extensions, Su et al. (2010) conclude that in empirical studies a higher level of sector disaggregation should be preferred. Lenzen (2011) addresses the mirror case in which the environmental data has a more detailed classification and studies whether the environmental data should be aggregated or the IO data should be disaggregated. Aggregating the environmental data implies an undesirable loss of detail; disaggregating the IO data can often only be done with incomplete data, which increases uncertainty about the validity of the final data set. Monte Carlo simulations indicate that disaggregating the IO data is preferable over aggregating the environmental data.

In a study of spatial aggregation, Su and Ang (2010) use an interregional IO model with eight regions for China. They find that the CO_2 emissions embodied in China's exports fall by 14% when the number of regions increases. Andrew et al. (2009) report that a world-average IO table offers a suitable substitute for an aggregate rest of the world (RoW) table, and they denounce single-country IO models — even though they admit it is better than ignoring imports altogether.

To extend and unify these efforts, we develop a second additive decomposition, now for sectoral and spatial aggregation errors.[3] Subsequently, we apply it to the measurement of embodied CO_2 emissions and embodied water use with the EXIOPOL database (Tukker et al., 2009), which offers more sectoral and environmental detail than any of the earlier mentioned studies or databases. In contrast to some earlier research, we show that aggregation errors are unacceptably large, especially for water use. We also conclude that outsourcing, contrary to popular belief, reduces instead of increases global warming and water use.

3.2 Methodology

3.2.1 The international input–output data used and the model built on it

The EXIOPOL database consists of an environmentally extended, international supply and use table (ISUT) for 2000, which includes all countries of the EU27, as well as 16 of the largest remaining countries (see Table 1.1). The EXIOPOL ISUT consists of a fully integrated intercountry use table in basic prices of the producing countries, including 43 submatrices for domestic use and 42×43 submatrices for bilateral trade flows (see Bouwmeester and Oosterhaven, 2009, for its construction methodology).[4] Each submatrix contains 129 product rows, with 129 industry and 5 final demand columns. The EXIOPOL ISUT also contains 43 domestic supply tables, each with 129 industry rows and 129 product columns (see Appendix A-5). Compared with the 59 × 59 detail of the ESA-95 classification, this ISUT disaggregates products and industries that are important from an environmental point of view, such as food and agricultural products, metal ores and products, mineral products, and energy products (Wood et al., 2010). In addition to the ISUT, the EXIOPOL database includes IO tables directly derived from the ISUT.

[3] The different types of aggregation errors may also interact, as noted by Su and Ang (2012).

[4] Trade flows with the RoW are accounted for in the database, but are not included in the model presented in this paper.

The EXIOPOL database also includes a wide range of environmental extensions (see Lutter et al., 2010). We focus on CO_2 emissions and water use, which present unique sectoral and spatial patterns.[5] Moreover, the former is an output of, whereas the latter is an input to the economic system. Most studies that compare production- with consumption-responsibility estimates look at CO_2 emissions (see Wiedmann, 2009, for an overview). In contrast, there is only a modest number of studies that investigate water use from a consumer perspective or consider the related issue of virtual water trade (Guan and Hubacek, 2007; Hoekstra and Chapagain, 2007; Yu et al., 2010; Hoekstra and Mekonnen, 2012). We aim to compare these results.

We use the EXIOPOL database to specify an international IO model that takes the international use of products, by industry and by final demand category, from the intercountry use table, and links this use of products to the output by industry by means of the market shares of products, by industry, from the domestic supply tables of the ISUT. This international industry-by-industry IO model thus features the following dimensions, shown as superscripts and subscripts of the variable at hand:

i, j	origin/destination industry subscripts running from 1 to I;
q	final demand category subscripts running from 1 to Q;
r, s	origin/destination country superscripts running from 1 to R; and
$\bullet$	summation over the index at hand.

All bilateral intermediate trade flows are specified by four indices: countries of origin and destination, and industries of origin and destination (Isard, 1951). Equation 3.1 gives a matrix representation of the IO model.[6]

The vectors $\mathbf{x}^r$ represent the total output vector of countries r. The diagonal submatrices $\mathbf{A}^{rr}$ show the domestic intermediate input coefficient matrices. The off-diagonal matrices $\mathbf{A}^{rs}$ show the imports coming from country r and used per unit of output of the industries of country s. The final demand vectors $\mathbf{f}^{r\bullet}$ reflect the country of origin, that is, they show the final demand of all countries for products produced in country r.

[5] Total water use in agricultural industries is the sum of the use of blue and green water. Blue water refers to rainwater evaporation; green water is ground- and surface water evaporation due to production. Agricultural water use estimates originate from the LPJmL model (Bondeau et al., 2007; Rost et al., 2008). Water use data for livestock and manufacturing industries has been modeled with the WaterGAP2 model (Alcamo et al., 2003).

[6] We denote matrices by bold capital letters, vectors by bold lower case letters, and scalars by italicized lower case letters. A prime indicates a transposed matrix or vector. A hat indicates a diagonal matrix of this vector. The vector $\mathbf{i}$ is a summation vector with ones. The identity matrix $\mathbf{I}$ thus equals $\hat{\mathbf{i}}$.

$$\begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^R \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} & \cdots & \mathbf{A}^{1R} \\ \mathbf{A}^{21} & \mathbf{A}^{22} & \cdots & \mathbf{A}^{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}^{R1} & \mathbf{A}^{R2} & \cdots & \mathbf{A}^{RR} \end{bmatrix} \begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^R \end{bmatrix} + \begin{bmatrix} \mathbf{f}^{1\bullet} \\ \mathbf{f}^{2\bullet} \\ \vdots \\ \mathbf{f}^{R\bullet} \end{bmatrix} \quad \textbf{(3.1)}$$

The EXIOPOL database also indicates the categorical and geographical destination of final demand. Each of the R final demand vectors in Equation 3.1 thus results from the following aggregation:

$$\mathbf{f}^{r\bullet} = \mathbf{F}^{r1}\,\mathbf{i} + \mathbf{F}^{r2}\,\mathbf{i} + \cdots + \mathbf{F}^{rR}\,\mathbf{i} + \mathbf{f}^{rRoW} \quad \textbf{(3.2)}$$

where $\mathbf{F}^{rs}$ represents an $I \times Q$ matrix of deliveries of products i from country r to the domestic final demand category q of country s. The domestic final demand categories include household consumption expenditures, government consumption expenditures, gross fixed capital formation, and changes in inventories and valuables. The last vector $\mathbf{f}^{rRoW}$ represents the exports of country r to the rest of the world (RoW).

If we omit the country superscripts, Equations 3.1–3.2 can be represented simply as $\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{f}$, which has the well-known solution $\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{f} = \mathbf{L}\mathbf{f}$ (e.g., Miller and Blair, 2009). The matrix $(\mathbf{I} - \mathbf{A})^{-1}$ is commonly referred to as the Leontief inverse and denoted by $\mathbf{L}$.

3.2.2 The environmental extension of the international IO model

Any ISUT can be extended with satellite accounts with additional information. These accounts often relate to the column totals of the use table, and represent variables such as labor used, water used, or CO_2 emitted by industry, by country (e.g., de Haan and Keuning 1996). Dividing, say, industry emissions by total industry output gives an $I \times R$ row $\mathbf{d}'$ with emission coefficients, which indicate the direct emissions per unit of output of each industry in each country. Multiplication with the total output by industry, by country, reproduces the direct emissions:

$$\mathbf{e}' = \mathbf{d}'\,\hat{\mathbf{x}} \quad \textbf{(3.3)}$$

where $\mathbf{e}'$ is the $I \times R$ row vector of pollutants emitted or material resources used by industry, by country. Total emissions or total material use at the world level, that is, for all 43 countries and 129 industries, then can be derived substituting the solution of Equation 3.1 in Equation 3.3, i.e., $e = \mathbf{d}'\mathbf{L}\mathbf{f}$.

Using Equation 3.2, we can disaggregate total e according to different points of view. To demonstrate how, we write it in its fullest possible partitioned form:

$$\begin{aligned}\mathbf{E} &= \begin{bmatrix} \mathbf{E}^{11} & \cdots & \mathbf{E}^{1R} & \mathrm{e}^{1RoW} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{E}^{R1} & \cdots & \mathbf{E}^{RR} & \mathrm{e}^{RRoW} \end{bmatrix} \\ &= \begin{bmatrix} \hat{\mathbf{d}}^{1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{\mathbf{d}}^{R} \end{bmatrix} \begin{bmatrix} \mathbf{L}^{11} & \cdots & \mathbf{L}^{1R} \\ \vdots & \ddots & \vdots \\ \mathbf{L}^{R1} & \cdots & \mathbf{L}^{RR} \end{bmatrix} \begin{bmatrix} \mathbf{F}^{11} & \cdots & \mathbf{F}^{1R} & \mathbf{f}^{1RoW} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{F}^{R1} & \cdots & \mathbf{F}^{RR} & \mathbf{f}^{RRoW} \end{bmatrix}\end{aligned} \qquad \textbf{(3.4)}$$

Note that the submatrices $\mathbf{E}^{rs} = \hat{\mathbf{d}}^{r} \sum_{k} \mathbf{L}^{rk}\mathbf{F}^{ks}$ have dimension $I \times Q$. They indicate which part of the direct emissions by industry i in country r can be explained by, i.e. attributed to, the domestic final demand of category q of country s.

Thus, matrix **E** combines all possible causes (types of final demand by country) with all possible impacts (direct emissions or resource use by industry, by country). Aggregation along a row of **E** gives the total emissions by industry, by country, which is the focus of the producer-responsibility principle. Aggregation along the columns of **E** gives the total worldwide emissions caused by the final demand of each category in each country, which is the focus of the extended consumer-responsibility principle. We use the word ‘extended’, in the sense that some part of worldwide emissions must be allocated to government expenditures and private investments.

3.2.3 Additive specification errors

With this full information, extended, international IO model, we study the specification errors that occur in a classical, single-country, consumer-responsibility study. Such a single-country study typically adds the foreign import coefficient matrix, which in our case equals the sum of 43 import coefficient matrices ($\sum_{s \neq r} \mathbf{A}^{sr}$), to the domestic input coefficient matrix $\mathbf{A}^{rr}$, while the associated Leontief inverse $(\mathbf{I} - \sum_{s} \mathbf{A}^{sr})^{-1}$ gets pre-multiplied with the row of domestic emission coefficients $\mathbf{d}^{r\prime}$ (e.g., Lenzen 1998; Wyckoff and Roop 1994).

Compared with the full model in Equation 3.4, we note two specification errors: the use of domestic emission coefficients, where foreign emission coefficients should have been used, and the use of a single-country Leontief inverse with domestic technology coefficients, where the full international Leontief inverse should have been used. In the literature, the total specification error is known as the domestic technology assumption (DTA) error. We look at the two underlying causes for this error, separately, and define them in such a way that their sum equals the total DTA specification error.

3.2.3.1 Error of using domestic emission coefficients

To get a pure measure of the partial error of using domestic emission coefficients we compare the outcomes of the extended, international IO model with the outcomes of the same model with domestic emission coefficients substituted for the foreign coefficients for all countries. From a practitioner's point of view, this calculation may seem strange, but the alternative decomposition is impossible, because it would require the combination of the complete set of foreign emission coefficients with a single-country IO model.[7] The *absolute* errors, by country of final demand, are thus measured as follows:

$$\tilde{e}^r - e^r = \begin{bmatrix} \mathbf{d}^r \\ \mathbf{d}^r \\ \vdots \\ \mathbf{d}^r \end{bmatrix}' \begin{bmatrix} \mathbf{L}^{11} & \mathbf{L}^{12} & \cdots & \mathbf{L}^{1R} \\ \mathbf{L}^{21} & \mathbf{L}^{22} & \cdots & \mathbf{L}^{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{L}^{R1} & \mathbf{L}^{R2} & \cdots & \mathbf{L}^{RR} \end{bmatrix} \begin{bmatrix} \mathbf{f}^{1r} \\ \mathbf{f}^{2r} \\ \vdots \\ \mathbf{f}^{Rr} \end{bmatrix} - \begin{bmatrix} \mathbf{d}^1 \\ \mathbf{d}^2 \\ \vdots \\ \mathbf{d}^R \end{bmatrix}' \begin{bmatrix} \mathbf{L}^{11} & \mathbf{L}^{12} & \cdots & \mathbf{L}^{1R} \\ \mathbf{L}^{21} & \mathbf{L}^{22} & \cdots & \mathbf{L}^{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{L}^{R1} & \mathbf{L}^{R2} & \cdots & \mathbf{L}^{RR} \end{bmatrix} \begin{bmatrix} \mathbf{f}^{1r} \\ \mathbf{f}^{2r} \\ \vdots \\ \mathbf{f}^{Rr} \end{bmatrix} \quad \textbf{(3.5)}$$

The last term of Equation 3.5 indicates the 'true' total worldwide emission caused by country *r*'s domestic final demand, while the first term indicates the estimate of this emission, made by assuming that the foreign $\mathbf{d}^s$ equals the domestic $\mathbf{d}^r$.[8]

In addition to absolute errors, which are of course larger for larger countries, we also study the related *relative* errors, which equal, in shorthand:

[7] In other words, our measure represents the only possible decomposition of the errors defined here. Hence, we do not average over multiple decompositions as is generally done when alternative decompositions are possible (see e.g. Dietzenbacher and Los, 1998).

[8] We use 'true' to indicate that this estimate is based on the most detailed set of data, which serves as our reference case, but which, of course, still contains unknown other errors.

$$\left(\tilde{e}^r - e^r\right) / e^r = \left(\mathbf{d}^{r\prime}\,\mathbf{L}\,\mathbf{f}^r - \mathbf{d}'\,\mathbf{L}\,\mathbf{f}^r\right) / \mathbf{d}'\mathbf{L}\,\mathbf{f}^r \tag{3.6}$$

with $\mathbf{f}^r$ defined as:

$$\mathbf{f}^r = \begin{bmatrix} \mathbf{f}^{1r} & \mathbf{f}^{2r} & \cdots & \mathbf{f}^{Rr} \end{bmatrix}' \tag{3.7}$$

This vector shows country r's consumption of final products of both domestic and foreign origin consumption goods and services. Therefore, $\mathbf{f}^r$ in Equation 3.7 defines final demand by country of destination, i.e., it reflects the extended consumer responsibility principle, whereas $\mathbf{f}^{r\bullet}$ in Equation 3.2 defines final demand by country of origin (i.e., country of production). The vector $\mathbf{f}^r$ in Equation 3.7 does not include the imports from the RoW, as R only represents the 43 individual countries for which data are provided in the EXIOPOL database. In addition, it does not include changes in domestic and foreign inventories and valuables, as this specific final demand category is a residual that does not represent actual demand. The *total relative world* error then equals:

$$\tilde{\varepsilon} = \left(\sum_r \tilde{e}^r - \sum_r e^r\right) / \sum_r e^r \tag{3.8}$$

We also attribute this error to *industry-specific* final demand, by means of:

$$\tilde{\varepsilon}_i = \left(\sum_r \tilde{e}_i^r - \sum_r e_i^r\right) / \sum_r e_i^r \tag{3.9}$$

where $e_i^r = \mathbf{d}'\mathbf{l}_i^r f_i^r$ and $\tilde{e}_i^r = \mathbf{d}^{r\prime}\mathbf{l}_i^r f_i^r$. Finally, we attribute the errors to *category-specific* final demand, with $\tilde{\varepsilon}_q$ calculated analogously to Equation 3.9.

3.2.3.2 *Error of using domestic technology coefficients*

To get a pure estimate of the partial error that results from using a single-country IO model with domestic technology coefficients, as opposed to an international IO model, we have to use the domestic emission coefficients of the country at hand in the both cases. Due to the nature of our problem, we cannot specify the first terms of Equation 3.10 with international emission coefficients. The *absolute* errors, by country of final demand, are calculated as follows:

$$\bar{e}^r - \tilde{e}^r = \mathbf{d}^{r\prime}\left(\mathbf{I} - \mathbf{A}^{\bullet r}\right)^{-1}\mathbf{f}^{\bullet r} - \mathbf{d}^{r\prime}\left(\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{f}^{r} \tag{3.10}$$

In the first term of Equation 3.10, we use $\mathbf{A}^{\bullet r} = \mathbf{A}^{lr} + \cdots + \mathbf{A}^{Rr}$ and $\mathbf{f}^{\bullet r} = \mathbf{f}^{lr} + \cdots + \mathbf{f}^{Rr}$,[9] whereas the last term equals the first term of Equation 3.5. Again, $\mathbf{f}^{\bullet r}$ in Equation 3.10 is not equal to either $\mathbf{f}^{r\bullet}$ in Equation 3.2 or $\mathbf{f}^{r}$ in Equation 3.7.

The absolute errors vary mainly by country size, therefore we also present *relative* errors, calculated as follows:

$$\left(\bar{e}^r - \tilde{e}^r\right)\Big/e^r = \left(\mathbf{d}^{r\prime}\left(\mathbf{I} - \mathbf{A}^{\bullet r}\right)^{-1}\mathbf{f}^{\bullet r} - \mathbf{d}^{r\prime}\,\mathbf{L}\,\mathbf{f}^{r}\right)\Big/\mathbf{d}^{\prime}\mathbf{L}\,\mathbf{f}^{r} \tag{3.11}$$

These relative errors are expressed as a ratio of the 'true' value e^r, instead of $\tilde{e}^r$. This choice is important, because only in this way can we add the first set of relative errors from Equation 3.6 to the second set of relative errors in Equation 3.11 and thereby obtain the *total* relative error of the classical single-country model compared with the full information, international IO model.

The relative domestic technology error for worldwide total emissions is defined as:

$$\bar{\varepsilon} = \left(\sum\nolimits_r \bar{e}^r - \sum\nolimits_r \tilde{e}^r\right)\Big/\sum\nolimits_r e^r \tag{3.12}$$

We again also present a disaggregation by causing *industry-specific* final demand analogous to Equation 3.9:

$$\bar{\varepsilon}_i = \left(\sum\nolimits_r \bar{e}_i^r - \sum\nolimits_r \tilde{e}_i^r\right)\Big/\sum\nolimits_r e_i^r \tag{3.13}$$

Finally, in all cases, the *total* relative specification error can be decomposed into specification errors due to using domestic technology coefficients and those due to using domestic emission coefficients:

[9] The total coefficient matrix $\mathbf{A}^{\bullet r}$, used here, deviates from the IO data published by the individual countries due to the estimation method of the international IO table (see Bouwmeester and Oosterhaven, 2009). The country-by-country trade flows have been made consistent by revaluing them in basic prices of the producing (i.e., exporting instead of importing) country.

$$\varepsilon = \frac{\sum_r \bar{e}^r - \sum_r e^r}{\sum_r e^r} = \frac{\sum_r \bar{e}^r - \sum_r \tilde{e}^r}{\sum_r e^r} + \frac{\sum_r \tilde{e}^r - \sum_r e^r}{\sum_r e^r} \tag{3.14}$$

3.2.4 Additive aggregation errors

We also investigate aggregation errors. In an international IO model, aggregation errors might be due to sectoral or spatial aggregation. The issue of aggregation bias has been subject of discussion for decades (e.g., Kymn 1990). Miller and Blair (2009) offer a recent overview that shows that the aggregation bias is more severe when the aggregated industries have very different structures.

To understand the nature of aggregation errors, it is important to be aware that, for any base year, Equation 3.4 will always result in exactly the same estimate of worldwide emissions, $e = \mathbf{i}'\mathbf{E}\mathbf{i}$, regardless of the actual level of sectoral or spatial aggregation. More specifically, the multiplication of any base year $\hat{\mathbf{d}}$ with $\mathbf{Lf}$, as in Equation 3.4, always reproduces e for that base year, irrespective of the aggregation level. Aggregation errors occur only when the structure of actual final demand differs from that in the base year.

To investigate the impact of the aggregation of industries and countries, we again evaluate the errors from the extended consumer-responsibility perspective. That is, we measure aggregation errors using the structural differences of the three main categories of domestic final demand in the EXIOPOL database, instead of using total final demand.

3.2.4.1 Sectoral aggregation errors

To measure the impact of the aggregation of industries, we calculate environmental footprints at three levels of sectoral detail. The 'true' value of the worldwide emissions embodied in final demand category q of country r is calculated at the most disaggregate level of 129 industries:

$$e_q^r = \begin{bmatrix} \mathbf{d}^1 \\ \mathbf{d}^2 \\ \vdots \\ \mathbf{d}^R \end{bmatrix}' \begin{bmatrix} \mathbf{L}^{11} & \mathbf{L}^{12} & \cdots & \mathbf{L}^{1R} \\ \mathbf{L}^{21} & \mathbf{L}^{22} & \cdots & \mathbf{L}^{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{L}^{R1} & \mathbf{L}^{R2} & \cdots & \mathbf{L}^{RR} \end{bmatrix} \begin{bmatrix} \mathbf{f}_q^{1r} \\ \mathbf{f}_q^{2r} \\ \vdots \\ \mathbf{f}_q^{Rr} \end{bmatrix} = \mathbf{d}'\mathbf{L}\mathbf{f}_q^r \tag{3.15}$$

This matrix results in $Q \times R = 129$ different 'true' values. In addition, Equation 3.15 is calculated at the level of the 59 industries of the EU27 IO tables, as published by Eurostat, and at the level of 10 aggregate sectors (see Appendix 3.1), which results in $e_q^r(59)$ and $e_q^r(10)$, respectively.

We compare only *relative* errors, sequentially, such that the *total* aggregation error equals the sum of the two separate errors:

$$\left[e_q^r(10) - e_q^r\right]/e_q^r = \left[e_q^r(10) - e_q^r(59)\right]/e_q^r + \left[e_q^r(59) - e_q^r\right]/e_q^r \tag{3.16}$$

The two separate errors do not necessarily have the same sign but may compensate for each other, such that the total error, even absolutely, can be smaller than each of the two partial errors separately.

The results of the comparison in Equation 3.16 will be presented at two levels. First, we discuss the sectoral aggregation errors for each of the three domestic final demand categories at the level of the 10 aggregate sectors, to determine which sectors are most heavily impacted. Second, we discuss the sectoral aggregation errors for each of the 43 countries to reveal which countries are most heavily impacted. For this last purpose, we post-multiply with *total* final demand by country, instead of with the separate categories, as in Equation 3.15:

$$e^r = \mathbf{d}'\mathbf{L}\mathbf{f}^r. \tag{3.17}$$

3.2.4.2 *Spatial aggregation errors*

We also measure the impact of spatial aggregation. For each country r, the spatial aggregation relates to a different set of the 42 remaining countries, $s \neq r$. The 'true' values of embodied worldwide water use and CO_2 emissions are calculated using the full international IO model with all 43 individual countries included, as in Equation 3.17. These 43 values then can be compared with calculations at two higher levels of aggregation for the remaining 42 countries:

1. The remaining countries are split into EU and non-EU countries, while both groups are further split into more developed and less developed countries. The latter subdivision is based on GDP/capita data (see Appendix 3.2), which results in an aggregation into five regions, where one of the remaining four regions is a little different for each country at hand.
2. The remaining countries are aggregated into one Rest of the World region. This aggregation into two regions results in a RoW that is a little different for each country.

The 3 × 43 outcomes for Equation 3.17 can be compared sequentially, as in Equation 3.16:

$$\left[e^r(2)-e^r\right]/e^r=\left[e^r(2)-e^r(5)\right]/e^r+\left[e^r(5)-e^r\right]/e^r, \qquad \textbf{(3.18)}$$

where e^r is worldwide embodied water use or CO_2 emissions allocated to total domestic final demand by country r, as calculated with the full 43-country model, and $e^r(2)$ and $e^r(5)$ are calculated with the aggregated models with two or five regions, respectively.

3.3 Results and discussion

In this section, we compare the differences in estimations of CO_2 and water use footprints derived from the alternative model specifications and varying levels of sectoral and spatial detail. The estimates using the full information, extended, international IO model with 43 countries and 129 industries/products are considered the ‘true’ values. We first discuss the specification errors that result from using domestic environmental and domestic interindustry coefficients. Then we turn to the differences stemming from sectoral and spatial aggregation.

3.3.1 Specification errors

Table 3.1 contains the specification errors according to Equation 3.9, calculated at the 129-industry level, but presented at the 59-industry level. The columns headed by $\tilde{e}-e$ represent the absolute error of using domestic environmental coefficients for foreign industries; the columns with $\tilde{\varepsilon}$ show the same error as a percentage of the ‘true’ value. Then the columns headed by $\bar{e}-\tilde{e}$ represent the absolute error of using domestic interindustry coefficients for foreign industries, and those headed by $\bar{\varepsilon}$ show the same error as percentage of the ‘true’ value. For each sector, the sum of the values in the columns $\tilde{e}-e$ and $\bar{e}-\tilde{e}$ reveals the value for the total error $\bar{e}-e$.

The row labeled 'Total net error' records the sum of the errors for columns with absolute numbers, but it provides the results of Equation 3.8 and Equation 3.12 for columns with percentage errors. Loosely defined, this information shows the total error at the world level, summed over all 43 countries. The 'Absolute deviations' row shows the sum of the absolute deviations for columns with absolute numbers and the weighted average of the absolute deviations for columns with percentage errors. This row provides a measure of the *gross* errors at the industry level. Finally, the last row of the table indicates the number of underestimations and overestimations.

Using domestic environmental coefficients leads to general overestimations for both embodied CO_2 emissions and embodied water use. The overestimation for CO_2 is 11087 Mt, or 63%, and for water it is 3164 km^3, or 18%. The unacceptably large overestimation of worldwide CO_2 partially cancels out when we use domestic technology coefficients, such that the overall net DTA (domestic technology assumption) error amounts to 37% (= 63% − 26%) of the full model estimate. For water, the overestimation is corrected only slightly to an overall net DTA error of is still 18% compared with the full model. The average absolute gross errors for water are only slightly higher than the total net errors. That is, though there are many sectors for which embodied water use is underestimated, the underestimations are not very large.

At the industry level, overestimation due to the use of domestic CO_2 coefficients occurs in all industries. In addition, the use of domestic interindustry coefficients reduces this overestimation in all industries. This also follows from the fact that the total net error equals the sum of the absolute errors. Hence, potentially high direct domestic emissions are reduced in all industries by outsourcing to countries with larger CO_2 coefficients.

The largest error resulting from the use of domestic CO_2 coefficients occurs in sector 23, coke, refined petroleum products and nuclear fuels. The absolute error for this sector amounts to 2566 Mt, almost one-quarter of the overall overestimation. In relative terms, the error is even more extreme and represents 977% of the 'true' estimate. In four other sectors, the relative error is also greater than 100%, and the maximum error in this set is 183% for sector 25, rubber and plastic products.

Table 3.1: Domestic environmental coefficients ($\tilde{\varepsilon}$) and domestic technology ($\bar{\varepsilon}$) specification errors, by industry

		CO_2 emissions				Water use			
		$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$	$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$
		Mt	%	Mt	%	km^3	%	km^3	%
i01	Agriculture, hunting	328	80	-204	-50	2149	31	-56	-1
i02	Forestry, logging	6	31	-2	-10	10	14	1	1
i05	Fishing	13	46	-7	-24	-14	-6	-2	-1
i10	Mining of coal and lignite - peat	4	19	-1	-5	0.002	0.1	-0.003	-0.2
i11	Crude petroleum and natural gas	25	59	-3	-6	-0.003	-0.02	-0.2	-2
i12	Mining of uranium and thorium ores	0.1	32	-0.2	-56	-0.2	-54	0.01	5
i13	Mining of metal ores	2	25	-1	-7	1	69	-0.3	-38
i14	Other mining and quarrying	2	36	-1	-10	0.2	22	-0.1	-10
i15	Food products and beverages	552	57	-259	-27	889	22	-132	-3
i16	Tobacco products	31	75	-17	-40	0.04	0.003	75	6
i17	Textiles	210	129	-176	-108	-26	-23	-15	-14
i18	Wearing apparel, fur	105	67	-94	-60	-23	-21	-15	-13
i19	Leather products	31	57	-27	-49	-9	-16	7	13
i20	Wood, cork and straw products	20	87	-5	-24	-5	-7	3	4
i21	Pulp, paper and paper products	53	94	-30	-53	-14	-11	13	10
i22	Publishing, printing, recorded media	84	82	-38	-37	-12	-20	-1	-1
i23	Coke, refined petroleum products, nuclear fuels	2566	977	-106	-40	-2	-14	1	7
i24	Chemicals and chemical products	559	138	-262	-64	15	11	-11	-8
i25	Rubber and plastic products	162	183	-135	-152	-38	-18	18	9
i26	Non-metallic mineral products	63	46	-44	-32	1	15	-2	-24
i27	Basic metals	54	34	-34	-21	13	160	-3	-36
i28	Fabricated metal products	107	55	-62	-32	18	131	-8	-55

Table 3.1: Domestic environmental coefficients ($\tilde{\varepsilon}$) and domestic technology ($\bar{\varepsilon}$) specification errors, by industry *(continued)*

		CO_2 emissions				Water use			
		$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$	$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$
		Mt	%	Mt	%	km^3	%	km^3	%
i29	Machinery and equipment n.e.c.	552	79	-313	-45	66	111	-37	-63
i30	Office machinery and computers	63	56	-38	-34	2	6	-12	-39
i31	Electrical machinery and apparatus n.e.c.	275	136	-200	-99	13	37	8	22
i32	Radio, television, communication equipment	176	81	-126	-58	-6	-14	-6	-15
i33	Medical, precision and optical instruments	71	72	-29	-29	9	59	-6	-40
i34	Motor vehicles, trailers and semi-trailers	612	86	-425	-60	20	20	-5	-5
i35	Other transport equipment	136	72	-91	-48	10	39	8	30
i36	Furniture; manufacturing n.e.c.	214	70	-191	-62	-17	-16	-8	-8
i37	Recycling*	-0.01	1	-0.04	3	0.1	29	-0.2	-70
i40	Electricity, gas, steam and hot water supply	149	7	-24	-1	0.1	1	-0.3	-1
i41	Collection, purification and distribution of water	10	22	-3	-7	-0.5	-14	-0.1	-4
i45	Construction	1058	35	-462	-15	-18	-2	14	2
i50	Sale, maintenance and repair of motor vehicles	91	48	-46	-24	-0.2	-0.3	1	2
i51	Wholesale trade and commission trade	132	59	-26	-12	2	3	-3	-5
i52	Retail trade; repair personal and household goods	219	37	-77	-13	13	6	0.3	0.1
i55	Hotels and restaurants	197	35	-96	-17	-21	-3	-41	-5
i60	Land transport; transport via pipelines	261	85	-102	-33	1	2	5	11
i61	Water transport	141	71	-144	-72	-0.3	-1	2	8
i62	Air transport	244	52	-86	-18	3	7	6	13
i63	Supporting transport activities; travel agencies	60	90	-9	-14	1	14	-1	-13
i64	Post and telecommunications	58	30	-21	-11	3	13	-4	-16
i65	Financial intermediation	35	37	-18	-19	23	100	-7	-32

Table 3.1: Domestic environmental coefficients ($\tilde{\varepsilon}$) and domestic technology ($\bar{\varepsilon}$) specification errors, by industry *(continued)*

		CO_2 emissions				Water use			
		$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$	$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$
		Mt	%	Mt	%	km^3	%	km^3	%
i66	Insurance and pension funding	35	44	-12	-14	7	34	0.03	0.1
i67	Activities auxiliary to financial intermediation	8	22	-5	-13	24	154	-1	-7
i70	Real estate activities	113	23	-33	-7	8	4	2	1
i71	Renting of machinery and equipment	7	31	-2	-8	0.01	0.3	0.1	2
i72	Computer and related activities	26	40	-11	-17	-1	-11	-1	-6
i73	Research and development	25	64	-9	-24	2	6	-1	-1
i74	Other business activities	36	36	-10	-10	-0.5	-1	1	2
i75	Public administration and defense	350	29	-141	-12	47	13	8	2
i80	Education	115	37	-17	-6	3	4	-1	-1
i85	Health and social work	429	65	-183	-28	20	6	15	4
i90	Sewage and refuse disposal, sanitation	20	61	-2	-5	0.3	3	1	6
i91	Activities of membership organization n.e.c.	31	27	-18	-16	-2	-5	1	2
i92	Recreational, cultural and sporting activities	72	40	-30	-17	-1	-2	-3	-4
i93	Other service activities	85	57	-62	-42	2	8	1	5
i95	Private households with employed persons	2	2	-2	-2	1	3	13	31
	Total net error	11087	63	-4570	-26	3164	18	-181	-1
	Absolute deviations	11087	63	4570	26	3586	21	585	3
	# of values <0 / >0*	1 / 58		59 / 0		22 / 37		33 / 26	

$\tilde{e}-e=$ absolute error of using domestic *environmental* coefficients for foreign industries; $\tilde{\varepsilon}=$ this error as % of the 'true' value; $\bar{e}-\tilde{e}=$ absolute error of using domestic *interindustry* coefficients for foreign industries; $\bar{\varepsilon}$ = this error as % of the 'true' value

* For i37, the CO_2 results show a sign reversion: CO_2 emissions in all three models are negative for this sector due to negative values for gross fixed capital formation. The full model value is more negative than the restricted model values.

Larger compensations of potentially high direct domestic emissions occur in industries with large errors due to using domestic CO_2 coefficients. Andrew et al. (2009) similarly report industries with significantly overestimated embodied emissions for the ten countries for which CO_2 emissions are most overestimated. Their study is based on GTAP data for 2001; they list petroleum and coal products and chemical, rubber and plastic products as significantly overestimated industries. However, the industries that are most often overestimated for the ten countries are oil extraction and gas extraction, which are not associated with large overestimations in our results.

The largest users of embodied *water* are i01, agriculture, hunting, and i15, food products and beverages. Both industries are associated with large absolute overestimations when using domestic water use coefficients. In terms of percentage errors, the metal industries (i27, i28 and i29) are noteworthy, as are financial intermediation services (i65 and i67), which both also use more water domestically than internationally, on average.

In Table 3.2 we provide the results of the specification errors for each of the 43 EXIOPOL countries. The total net errors are, of course, the same as those in Table 3.1; the value of the sum of both specification errors (i.e., DTA $= \tilde{e} - e + \bar{e} - \tilde{e}$) is what comparable studies generally restrict themselves to (e.g., Andrew et al., 2009).

Regarding embodied *CO_2 emissions* at the individual country level, using domestic interindustry coefficients partially cancels out the errors made by using domestic emission coefficients for 25 countries. In terms of the magnitude of the absolute errors, Germany stands out; it is responsible for 5201 (47%) of the total net error of 11087 Mt. If we add the *absolute* errors of India and the United States, we reach 85% of the total. Using domestic CO_2 coefficients severely overestimates these countries' emissions. However, in the single-country model, using domestic interindustry coefficients cancels out the impact of the high domestic CO_2 coefficients for both India and the United States. In terms of large *percentage* errors, Germany again stands out with 663%, joined by India (259%) and The Netherlands (284%).

Compared with the 2001 data of Andrew et al. (2009), we find the same sign (over- or underestimation) for the difference between the full model and the one with both domestic technology assumptions (DTA) for 26 of the 42 countries that the two data sets have in common. The average overestimation over all 42 countries is 31% in our work, versus 12% in the Andrew et al. (2009) study. The most extreme combined overestimation with Andrew et al. (2009) is 236% for Estonia; our most extreme overestimation is 643% (= 663% – 20%) for Germany. The underestimations reveal a larger difference between estimates in their work: –59% for Switzerland versus our result of –42% for Switzerland.

Table 3.2: Domestic environmental coefficients ($\tilde{\varepsilon}$) and domestic technology ($\bar{\varepsilon}$) specification errors, by country

	CO_2 emissions				Water use			
	$\tilde{e} - e$	$\tilde{\varepsilon}$	$\bar{e} - \tilde{e}$	$\bar{\varepsilon}$	$\tilde{e} - e$	$\tilde{\varepsilon}$	$\bar{e} - \tilde{e}$	$\bar{\varepsilon}$
	Mt	%	Mt	%	km^3	%	km^3	%
Australia	5	2	8	3	43	34	-40	-31
Austria	-5	-7	-11	-15	-8	-29	-1	-4
Belgium	-10	-12	9	11	-14	-44	-2	-7
Brazil	28	6	-14	-3	-102	-9	12	1
Bulgaria	4	15	6	19	1	3	1	5
Canada	35	11	-16	-5	100	34	-51	-18
China	671	23	-223	-7	-64	-4	35	2
Cyprus	4	52	1	20	-3	-64	0.3	8
Czech Republic	4	4	14	15	-3	-7	2	6
Denmark	-12	-28	-2	-4	-7	-42	-0.4	-2
Estonia	-1	-8	4	33	2	40	2	43
Finland	40	77	-44	-84	-4	-20	0.4	2
France	210	59	-77	-22	-50	-22	12	5
Germany	5201	663	-154	-20	96	37	-95	-37
Greece	21	21	-17	-16	-4	-7	2	4
Hungary	13	30	-3	-6	3	10	3	11
India	2408	259	-2104	-226	-35	-1	23	1
Indonesia	19	9	-11	-5	205	6	222	7
Ireland	-1	-2	-2	-4	-3	-19	1	4
Italy	57	14	-49	-12	-28	-15	9	5
Japan	-238	-16	-2	-0.1	-572	-82	-6	-1
Latvia	1	13	-0.3	-3	-0.1	-0.2	-1	-1
Lithuania	-2	-17	-2	-14	-8	-18	-0.3	-1
Luxembourg	-1	-30	-0.1	-3	-0.4	-17	1	26
Malta	1	25	2	54	-1	-99	0.001	0.2
Mexico	171	53	-69	-21	519	48	-80	-7
Netherlands	389	284	8	6	-37	-86	0.4	1
Norway	37	98	-21	-57	251	1175	-76	-357
Poland	33	12	28	11	0.1	0.03	-1	-1
Portugal	6	9	6	9	-2	-4	5	12
Romania	22	32	11	16	4	5	13	16
Russian Fed.	145	22	9	1	3229	270	-153	-13
Slovak Republic	9	34	1	4	0.3	1	1	4
Slovenia	-1	-11	1	5	-3	-52	1	12
South Africa	20	10	20	10	26	19	-5	-3
South Korea	-33	-8	24	5	-104	-61	-8	-5
Spain	53	20	-16	-6	-5	-2	-6	-3
Sweden	51	83	-56	-90	10	27	-12	-31
Switzerland	-23	-35	-5	-8	20	44	1	2
Taiwan	-16	-8	13	6	1	1	-33	-44

Table 3.2: Domestic environmental coefficients ($\tilde{\varepsilon}$) and domestic technology ($\bar{\varepsilon}$) specification errors, by country *(continued)*

	CO_2 emissions				Water use			
	$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$	$\tilde{e}-e$	$\tilde{\varepsilon}$	$\bar{e}-\tilde{e}$	$\bar{\varepsilon}$
	Mt	%	Mt	%	km^3	%	km^3	%
Turkey	-13	-6	8	4	7	7	-16	-15
UK	-59	-11	-19	-4	-80	-53	-10	-7
US	1845	35	-1825	-34	-216	-7	70	2
Total net error	11087	63	-4570	-26	3164	18	-181	-1
Absolute deviations	11918	68	4911	28	5868	34	1015	6
# of values <0 / >0	14 / 29		25 / 18		25 / 18		20 / 23	

$\tilde{e}-e$ = absolute error of using domestic *environmental* coefficients for foreign industries; $\tilde{\varepsilon}$ = this error as % of the 'true' value; $\bar{e}-\tilde{e}$ = absolute error of using domestic *interindustry* coefficients for foreign industries; $\bar{\varepsilon}$ = this error as % of the 'true' value

The specific year of the estimations also seems to matter, as similar work by these same authors for 2004 shows a different pattern of under- and overestimations (Peters, Andrew and Lennox 2011). Compared with this latter set, we find the same sign for only 22 of the 42 countries. However, Peters, Andrew and Lennox (2011) find even more dramatic overestimations from the use of the combined DTA compared with the estimates of their full model: 9875% for Mexico, 1567% for South Africa, 1371% for Russia, and 1098% for Greece.[10] The average difference between our results and Peters, Andrew and Lennox (2011) is 393%, whereas the average difference when we compared our results with the Andrew et al. (2009) was only 48%.

In the results for embodied *water use*, the errors associated with applying domestic water use coefficients range from –572 km^3 for Japan to +3229 km^3 for the Russian Federation. Water is particularly abundant in the Russian Federation, which is most likely the reason why it is abundantly used in production. This country alone accounts for 55% of the total of absolute errors caused by using domestic water use coefficients. The 3229 km^3 overestimation corresponds to 270% of the 'true' estimate, whereas the average absolute deviation amounts to 34%. The impact in percentage terms of using domestic water coefficients is even more remarkable for Norway, with a value of

[10] We present the differences between their DTA model and their MRIO 'Transport exogenous' model. The GTAP database includes an international transport pool, which records the product being transported but not the bilateral link between the supplier and user of the transport service. In the 'Transport exogenous' model, the pool is not allocated to use sectors, such that it gets treated as a final demand category, a designation that is adjusted in the 'Transport endogenous' model. However, the authors issue several caveats and warn that this model should not be used to calculate consumption-based emissions without additional verification and model development.

+1175%. This extremely large percentage actually represents the very low water use estimated in the full model (21 km^3), while the estimate based on domestic water coefficients still is not very large compared with other countries.

The use of domestic interindustry coefficients for embodied water use leads for six countries to relatively large absolute errors: Indonesia and the United States show an overestimation, whereas the Russian Federation, Germany, Mexico, and Norway show underestimations. Again, in percentage terms, the most remarkable country is Norway, with a relative error of –357%. Most countries show comparatively small absolute errors. Embodied water use in the countries of the European Union is underestimated in 19 of the 27 cases if we consider the combined DTA effect of using both domestic emission and domestic technology coefficients in a single-country model.

3.3.2 Aggregation errors

3.3.2.1 Sectoral aggregation

Table 3.3 contains the percentage differences in the estimated CO_2 and water use footprints when the sectoral detail is reduced from 129 to 59 to 10 sectors. The results are presented by sector (10-sector level) and by domestic final demand category.

The results for CO_2 emissions are rather mixed, except for the persistent underestimation of CO_2 emissions by construction (sector F). The largest errors, all overestimations, arise for government consumption and relate to the electricity/gas/water sector (E) and the financial intermediation/real estate/business services sector (JK). Most errors are relatively small and within a single-digit range. The largest errors occur when we aggregate further from 59 to 10 sectors.

The percentage errors resulting from the aggregation from 129 to 59 industries are far more prominent for embodied water though. Extreme percentage errors, mostly overestimations, occur in the aggregation from 59 to 10 sectors for water use. Sector E (Electricity, gas and water supply) shows the largest positive percentage errors, especially for household and government consumption. Water use percentage errors resulting from the aggregation to 10 sectors are relatively large and positive for all sectors when the specific structure of gross fixed capital formation is used for weighting.

Table 3.3: Sectoral aggregation errors, by final demand, by impacted sector, in percentages

	CO_2 emissions			Water use		
%	129 >> 59	59 >> 10	129 >> 10	129 >> 59	59 >> 10	129 >> 10
Final demand 1: household consumption						
ABC – Primary*	-0.4	19	19	9	-32	-22
D – Manufacturing	1	19	20	-1	-33	-34
E – Public utility	18	2	20	-46	2232	2186
F – Construction	-6	-27	-33	-51	220	169
GH – Trade	-2	0.3	-2	23	-40	-17
I – Transport	4	-11	-7	-16	121	106
JK – Financial	-3	18	15	-51	66	15
L – Government	1	-3	-3	-36	79	43
MN – Quaternary	-4	-0.4	-5	-26	28	2
OPQ – Other	-4	2	-2	-23	54	31
Final demand 2: government consumption						
ABC – Primary	-16	19	3	-2	-26	-28
D – Manufacturing	-4	8	4	-13	-0.4	-14
E – Public utility	-8	87	78	-32	1595	1563
F – Construction	-14	-30	-44	-48	131	83
GH – Trade	8	-10	-2	8	-46	-39
I – Transport	2	8	10	-13	114	101
JK – Financial	13	35	48	36	27	64
L – Government	2	-1	0.3	-28	61	32
MN – Quaternary	4	3	7	-17	50	34
OPQ – Other	-3	-1	-4	-16	4	-12
Final demand 3: gross fixed capital formation						
ABC – Primary	-3	-3	-6	12	111	123
D – Manufacturing	-1	11	11	-47	496	449
E – Public utility	13	6	20	-44	903	859
F – Construction	-5	-34	-39	-53	124	72
GH – Trade	-0.2	17	17	-37	147	109
I – Transport	13	-8	5	-3	121	119
JK – Financial	-5	-2	-7	-46	90	44
L – Government	-8	5	-3	-52	234	182
MN – Quaternary	8	6	14	5	526	531
OPQ – Other	-2	-4	-7	-21	35	14

* For a more detailed description of the sectors see Appendix 3.1

In Table 3.4 we report the sectoral aggregation errors by country. For *CO_2 emissions*, most aggregation errors are quite small. The larger errors are related to the aggregation to 10 sectors, but in that case, they almost cancel out at the world level. The smaller errors found when aggregating to 59 industries are mostly positive and cancel out less. The total net error of +1.9% confirms this finding; it is caused mostly by the aggregation from 129 to 59 industries. Two countries have sizeable total errors of 30% or more, whereas all other countries have errors less than 20%: Russia has a larger error when aggregating from 129 to 59 industries, whereas Norway's error is almost fully due to the aggregation from 59 to 10 sectors.

Su et al. (2010) suggest that a sector detail of around 40 industries is sufficient to capture the majority of CO_2 emissions embodied in production. However, they use a single-country model and only investigate emissions embodied in exports of China and Singapore. Our results support their finding, to the extent that smaller errors occur when aggregating from 129 to 59 industries. Although we find sizeable errors for aggregating further, they may occur at any point in the 59-to-10-sector interval. The results also can differ strongly across countries, suggesting that a uniform prescription for the level of sector detail is not possible.

Sectoral aggregation errors for embodied *water use* are much greater than those for CO_2 emissions. For most countries, water use is underestimated when aggregating from 129 to 59 industries. With a total net value of –1.3% the underestimation of the world average seems limited; however, the average absolute deviation of 9.2% indicates that most errors are not small. Five countries are associated with large overestimations, and Turkey is especially notable with an overestimation of water use by 79%. The percentage errors obtained from aggregating to 59 industries partially cancel out when aggregating further to 10 sectors, but at the country level, these errors appear unacceptably large. The largest overestimation of 87% occurs for Korea, and the largest underestimation, 54%, relates to Russia.

In terms of total net error, we also find for water use that the average world error percentage is mostly determined by the aggregation from 59 to 10 sectors. Again, we observe that aggregating from 129 to 59 industries results in smaller errors than aggregating further to 10 sectors, for both individual countries and the world as a whole. This finding seems remarkable when we consider that the 129-sector classification is especially detailed for agriculture and food, both of which use a lot of water.

Table 3.4: Sectoral aggregation errors, by country, in percentages

	CO_2 emissions			Water use		
%	129 >> 59	59 >> 10	129 >> 10	129 >> 59	59 >> 10	129 >> 10
Australia	-1	7	6	-17	8	-9
Austria	6	-2	3	-11	47	36
Belgium	2	5	7	-7	21	14
Brazil	1	6	7	-6	-6	-11
Bulgaria	7	13	20	-0.5	48	47
Canada	1	3	4	-18	-3	-21
China	3	-4	-1	1	-1	-0.4
Cyprus	-0.1	1	1	5	28	33
Czech Republic	3	1	5	-8	29	21
Denmark	5	11	16	-17	34	16
Estonia	3	-4	-1	7	-3	4
Finland	2	-1	1	-3	49	46
France	1	3	3	-3	18	15
Germany	2	4	6	-1	42	41
Greece	1	18	19	-5	29	24
Hungary	-1	-4	-5	-27	20	-8
India	-2	-0.2	-2	-1	-3	-4
Indonesia	-5	7	1	17	-32	-15
Ireland	-0.1	-16	-16	-12	23	12
Italy	0.3	-1	-1	-12	32	21
Japan	1	3	4	-39	68	28
Latvia	0.3	6	7	-46	-6	-51
Lithuania	3	0.4	4	3	-5	-2
Luxembourg	5	15	20	-35	39	3
Malta	2	-8	-7	39	12	51
Mexico	-6	-7	-13	-3	-31	-34
Netherlands	3	8	11	7	40	47
Norway	2	28	30	-23	26	3
Poland	2	0.3	3	-16	5	-11
Portugal	-1	-4	-5	0.4	10	10
Romania	3	8	11	-24	4	-20
Russian Fed.	24	13	36	3	-54	-51
Slovak Republic	4	10	14	-12	40	27
Slovenia	4	-5	-1	-24	27	2
South Africa	5	12	17	3	-49	-46
South Korea	-5	-0.5	-5	-41	87	46
Spain	1	0.3	1	-11	14	4
Sweden	1	4	5	-4	3	-1
Switzerland	0.5	-7	-7	-44	12	-32
Taiwan	-2	-5	-7	-10	61	52
Turkey	-0.3	-7	-7	79	14	93
U.K.	1	-2	-1	1	33	34
U.S.	0.1	-1	-1	-6	23	17
Total net error	1.5	0.4	1.9	-1.3	-2.7	-4.0
Average abs. dev.	2.3	3.4	4.4	9.2	23.2	18.1
# of values <0 / >0	11 / 32	18 / 25	16 / 27	31 / 12	11 / 32	16 / 27

3.3.2.2 Spatial aggregation

Table 3.5 contains the spatial aggregation errors. Each country's embodied CO_2 emissions and embodied water use from the full intercountry IO model is compared with estimations using an IO model with the specific country, combined with four aggregate regions: the poorer and the richer Rest of the EU, and the poorer and richer RoW. A further aggregation uses an IO model for the country at hand and one single, large RoW (see Appendix 3.2).

The absolute percentage deviation shows an average overestimation of 3.0% of the *CO_2 emissions* when aggregating to two regions. However, far more relevant for policy formulation and negotiation is that the total error for individual countries ranges from +38% for Luxembourg to –29% for Lithuania. The overestimation at the world level is most pronounced in the aggregation from five to two regions, with a value of 2.4%. These findings are in line with the overestimation found by Andrew et al. (2009) for OECD countries in models with a limited amount of regions included.

In terms of individual countries, when aggregating from 43 to 5 regions, most aggregation errors of CO_2 footprints are small, except those of the Baltic countries and Cyprus. In the Baltic States, embodied CO_2 emissions are underestimated by the five-region model, whereas aggregation in the Cyprus case leads to overestimation by 17%, which is fully compensated when the model is reduced further to two regions. In the aggregation from five to two regions, all large errors are overestimations of CO_2, except for the error for Cyprus. In general, at the individual country level, we observe that aggregation errors when going from 43 to 5 regions tend to be smaller than when aggregating further to 2 regions. Thus, a carefully designed aggregation, such as our four-region classification, may be acceptable. Blair and Miller (1983), investigating the aggregation of regions in a multiregional IO model, similarly conclude that questions pertaining to one or a few specific regions may be studied in a model in which the rest of the regions are aggregated into one 'rest of' region.

This optimistic conclusion is refuted for *water use* though. The errors of going from 43 to 5 regions show less extreme values than those of further aggregating to 2 regions, but both types of errors are unacceptably large. What is worse, they do not compensate for each other, as the last column of Table 3.5 shows. At the individual country level, Luxembourg and Switzerland stand out with overestimations of, respectively, 193% and 146% of their water use footprint when trading partners get more aggregated. Again, underestimation occurs generally when aggregating to four general regions, whereas overestimation occurs mostly when further aggregating to one large RoW region.

Table 3.5: Spatial aggregation errors, by country, in percentages

	CO_2 emissions			Water use		
%	43 >> 5	5 >> 2	43 >> 2	43 >> 5	5 >> 2	43 >> 2
Australia	0.1	2	2	-21	3	-18
Austria	-2	14	12	-1	54	53
Belgium	1	20	21	-15	106	91
Brazil	-1	1	-0.01	-7	-1	-8
Bulgaria	-3	-2	-5	4	-2	2
Canada	1	4	6	-15	11	-4
China	-1	0.2	-0.3	-3	-1	-4
Cyprus	17	-17	0.2	84	-12	72
Czech Republic	-3	2	-1	-3	11	8
Denmark	-3	14	11	-9	53	44
Estonia	-9	0.1	-9	-4	10	6
Finland	-5	5	0.01	-5	18	13
France	-0.4	16	16	-6	29	23
Germany	-1	7	6	-3	37	34
Greece	-1	4	3	-2	21	18
Hungary	-4	6	2	-4	8	4
India	-1	-1	-1	-1	-0.5	-1
Indonesia	-1	0.4	-1	0.03	-0.02	0.01
Ireland	2	17	18	-7	68	61
Italy	-1	7	6	-9	33	24
Japan	-3	-1	-3	-16	-3	-19
Latvia	-21	-7	-28	-8	1	-7
Lithuania	-22	-7	-29	-20	1	-19
Luxembourg	-0.03	38	38	-6	200	193
Malta	8	3	11	-13	101	89
Mexico	-3	5	2	-3	1	-2
Netherlands	-3	17	14	-13	63	50
Norway	-4	20	16	6	87	93
Poland	-2	3	0.3	-3	4	1
Portugal	2	14	16	-4	43	38
Romania	-2	1	-1	0.1	2	2
Russian Federation	-0.3	0.4	0.2	2	-1	2
Slovak Republic	-6	0.2	-6	-5	15	10
Slovenia	-0.1	14	14	-12	26	15
South Africa	-0.1	1	1	-4	1	-4
South Korea	-3	-0.3	-3	-27	-11	-38
Spain	2	9	11	-3	16	12
Sweden	2	22	23	-5	37	32
Switzerland	-0.3	26	26	29	117	146
Taiwan	-1	0.3	-1	-30	6	-24
Turkey	-6	1	-5	-7	8	0.4
United Kingdom	0.1	9	9	-14	45	31
United States	1	-0.2	1	-7	2	-4
Total net error	-0.4	2.1	1.7	-4.0	3.5	-0.5
Average. abs. dev.	1.4	2.4	3.0	4.5	4.5	6.4
# of values <0 / >0	32 / 11	8 / 35	15 / 28	36 / 7	9 / 34	13 / 30

3.4 Conclusion and contributions

As opposed to the standard practice of investigating the error of a combined domestic technology assumption (DTA), we have developed a methodology to estimate model specification errors and aggregation errors, such that the separate constituent errors consistently sum to the total DTA error.

The partial *model specification* errors relate to the use of domestic emission coefficients for foreign industries and of a single-country IO framework with domestic interindustry coefficients instead of an intercountry one. The empirical outcomes show that for CO_2 footprints, using domestic emission coefficients grossly overestimates the embodied emissions. However, most large errors (partially) cancel out when further simplifying the specification with the assumption that domestic interindustry coefficients hold for foreign industries. This effect is even more prevalent in the industry-by-industry results, in particular for service industries.

These findings indicate that emissions coefficients and production structure are closely interrelated. Most countries with relatively high emission coefficients tend to reduce their worldwide CO_2 footprint by outsourcing to countries with lower emission coefficients. This contradicts the popular belief that outsourcing increases global warming.

For embodied water use, the range of the two specification errors is wider than that for CO_2 emissions, whereas they hardly compensate for each other, as opposed to the partial errors in the CO_2 estimates. The deviation from the full model estimates is caused primarily by applying domestic water use coefficients to foreign industries. At the individual country level, Norway stands out for its large relative errors in the estimation of embodied water use.

The *sectoral aggregation* errors are evaluated for aggregating the 129 EXIOPOL industries to the 59 EU industries and then further to 10 aggregate sectors. The empirical outcomes show that this aggregation has a much larger impact on the estimation of embodied water use than on that of CO_2 emissions. In addition, larger errors occur in the aggregation from 59 to 10 sectors than in that from 129 to 59 industries, which is remarkable since the 129-industry level was especially designed to capture differences in environmental coefficients better.

The finding that the aggregation from 59 to 10 sectors results in higher errors than the aggregation from 129 to 59 is in line with prior findings, such as Lenzen et al. (2004) for Denmark and Su et al. (2010) for China. However, the difference we find between the errors for embodied water use versus those for CO_2 footprints limit the usefulness of any specific recommendation related to the overall number of industries to be included. The focus of the specific study and the prevailing differences in production structures, when using a more detailed sector classification, should be the driving arguments in the determination of the required level of sectoral detail.

Spatial aggregation errors are evaluated for aggregating the remaining 42 countries of each of the 43 EXIOPOL countries to four broad regions, defined as inside or outside the EU, and as relatively poor or rich, and then further aggregating these four regions to a single large 'rest of the world' region. The results for CO_2 footprints suggest that a carefully designed spatial aggregation, such as the first one, is acceptable. Yet this provisional conclusion is immediately refuted by the more extreme results for embodied water use. Aggregating the remaining 42 countries into four main regions leads to a considerable underestimation of the water use footprint of many countries. Further aggregation to one individual country and a RoW region leads to even larger errors.

Thus, the effects of different aggregation options are more sizeable for water use than for CO_2 emissions. The likely explanation is that only a few industries are linked to large water use coefficients, whereas each and every sector uses energy and therefore produces CO_2 emissions, though with different intensities. The mixed effects of aggregation show that a universal optimal level of sector detail cannot be readily identified. Focusing on disaggregating sectors, only to provide more detail, may actually be suboptimal compared to improving the quality of currently available data.

Overall, this study has shown that differences in specification or aggregation can have sizeable effects on the estimates of embodied emissions or resource use. Unfortunately, no general recipe is available for assessing environmental issues with an input–output framework. It is really up to the analyst to argue, in each specific application, why the chosen specification and the chosen level of detail are justified. By comparing CO_2 and water use, we reveal that using a more simple representation may be more justifiable in cases in which environmental footprints are caused by all industries and all countries. When a relatively limited number of sectors, or countries, is responsible for the environmental footprint, it appears the more elaborate representation of those sectors and countries is preferable.

Appendix 3.1: Sector aggregation

	NACE Rev. 1.1 classification*	EXIOPOL industries				
A	Agriculture, hunting and forestry	i01	i02			
B	Fishing	i05				
C	Mining and quarrying	i10	i11	i12	i13	i14
D	Manufacturing	i15	i16	i17	i18	i19
		i20	i21	i22	i23	i24
		i25	i26	i27	i28	i29
		i30	i31	i32	i33	i34
		i35	i36	i37		
E	Electricity, gas and water supply	i40	i41			
F	Construction	i45				
G	Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and	i50	i51	i52		
H	Hotels and restaurants	i55				
I	Transport, storage and communication	i60	i61	i62	i63	i64
J	Financial intermediation	65	i66	i67		
K	Real estate, renting and business activities	i70	i71	i72	i73	i74
L	Public administration and defense; compulsory social security	i75				
M	Education	i80				
N	Health and social work	i85				
O	Other community, social and personal service activities	i90	i91	i92	i93	
P	Activities of households	i95	i96	i97		
Q	Extra-territorial organizations and bodies	i99				

* A description of the NACE Rev. 1.1 classification is available at: http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=DSP_GEN_DESC_VIEW_NOHDR&StrNom=NACE_1_1&StrLanguageCode=EN

Appendix 3.2: Spatial aggregation

EU-high	PPP $ 2000*
Luxembourg	53652
Netherlands	29403
Denmark	28829
Austria	28773
Ireland	28639
Sweden	27961
Belgium	27612
United Kingdom	26072
Germany	25945
Finland	25653
Italy	25595
France	25328
Spain	21323

EU-low	PPP $ 2000*
Cyprus	19412
Greece	18412
Malta	18291
Portugal	17751
Slovenia	17474
Czech Republic	14993
Hungary	12266
Slovak Republic	10997
Poland	10514
Estonia	9882
Lithuania	8602
Latvia	8031
Bulgaria	6301
Romania	5654

Other-high	PPP $ 2000*
Norway	36130
United States	35081
Switzerland	31731
Canada	28407
Australia	26422
Japan	25619
Taiwan	19866
Korea, Rep.	17219

Other-low	PPP $ 2000*
Mexico	9201
Turkey	8867
Brazil	7021
Russian Federation	6824
South Africa	6773
Indonesia	2417
China	2364
India	1574

* GDP/capita, in PPP (current international $) for 2000. World Development Indicators, http://data.worldbank.org/indicator, accessed 25-3-2011. Source: World Bank national accounts data, and OECD National Accounts data files.